APPENDIX A

EVIDENCE	INVOCATION OF PRIVILEGE	SIGNIFICANCE	DATE OF COMMUNICATION
Exhibit A, Meta_Kadrey_00057872 at 57872-73	This document, entitled "Datasets Considerations" and labeled "A/C PRIV," contains heavy redactions over discussions of certain datasets that "will benefit the eng[ineering] team if possible to use." It explicitly mentions books and Libgen, and wanting to get "as much long form writing as possible in the next 4-6 weeks[.]"	These redactions show attorney-client communications in furtherance of Meta's decision to obtain and use massive quantities of copyrighted material that it knew to be pirated, and show that Meta believed there was time pressure to obtain this material. This document is dated approximately one week before Meta decided to stop its efforts to license copyrighted works. *See* Ex. F. Meta's 30(b)(6) witness and outside counsel claimed privilege over "considerations" in Meta's decision to stop licensing efforts even after Plaintiffs successfully moved to compel a further 30(b)(6) deposition based on such privilege assertions. *See* Ex. G.	April 1, 2023
Exhibit B, Meta_Kadrey_00088838 at 88838-41	This message chain discusses needing "zuck/cox/ahmad" approval for the use of certain datasets including Books3. It appears to convey (redacted) legal advice from "ahuva's skip," presumably referencing legal advice from a senior lawyer.	These redactions show attorney-client communications in furtherance of Meta's decision to use copyrighted data it knew to be pirated. These messages are dated the same day (or the day before) Meta decided to stop pursuing licenses for copyrighted works. *See* Ex. F. Meta's 30(b)(6) witness and outside counsel claimed privilege over "considerations" in Meta's decision to stop licensing efforts even after Plaintiffs successfully moved to compel a further 30(b)(6) deposition based on such privilege assertions. *See* Ex. G.	April 6, 2023

EVIDENCE	INVOCATION OF PRIVILEGE	SIGNIFICANCE	DATE OF COMMUNICATION
Exhibit C, Meta_Kadrey_00048977	Comments from Ahuva Goldstand (meta in-house counsel), Dustin Holland, and Melanie Kambadur are redacted from a document discussing "approval to leverage 3P [third party] pre-training dataset(s) for the purpose of further training the Llama model[.]" The section on which the comments are embedded mentions "high risk IP" and the section immediately preceding the comments lists the datasets needing approval, which includes pirated database Books3. This document contains hyperlinks to two documents at Meta_Kadrey_00048977, titled "a/c priv - Third-Party GenAl LLM Training Dataset Pll & Scraping Principles" and "LLM Safety: PII Mitigations." These documents appear to have been fully withheld, and Plaintiffs cannot ascertain from Meta's privilege logs any additional information, but the title of the first one at least indicates it concerns datasets Meta obtained from third-party shadow libraries.	These redactions show attorney-client communications in furtherance of Meta's April 2023 decision to stop efforts to license copyrighted material and instead take (using peer-to-peer networks) and use massive quantities of copyrighted works it knew to be pirated and "high risk."	May 11–15, 2023

EVIDENCE	INVOCATION OF PRIVILEGE	SIGNIFICANCE	DATE OF COMMUNICATION
Exhibit D, Meta_Kadrey_00054898	A conversation between Melanie Kambadur and Nisha Deo, apparently discussing responding to a press inquiry from The Atlantic, is largely redacted. The questions include: • "Was Meta aware at the time that Books3 includes pirated books?" • "Was Meta aware at the time that Books3 includes books that have been stripped of copyright information, possibly in violation of the DMCA?"	These redactions show attorney-client communications in furtherance of Meta's decision to stop efforts to license copyrighted works and instead take (using peer-to-peer networks) and use copyrighted works that it knew to be pirated and involvement in lying to the press about Meta's use of said data.	August 18, 2023
Exhibit E, Meta_Kadrey_00079062 at 79063-64	This document discusses LLM model outputs regurgitating copyrighted material. Meta employees appear to seek guidelines or requirements for reporting copyrighted output. The responses are redacted.	These redactions show attorney-client communications in furtherance of Meta's decision to take (using peer-to-peer networks) and use copyrighted data it knew to be pirated and show that Meta in-house counsel was aware that Llama models regurgitated copyrighted material.	April 15, 2024